Protein Engineering, Design & Selection vol. 17 no. 9 pp. 699–707, 2004
Published online November 2, 2004 doi:10.1093/protein/gzh082

Covalent DNA display as a novel tool for directed evolution of proteins *in vitro*

Julian Bertschinger[1] and Dario Neri

Institute of Pharmaceutical Sciences, ETH Hönggerberg,
Wolfgang-Pauli-Strasse 10, HCI G394, 8093 Zurich, Switzerland

[1]To whom correspondence should be addressed.
E-mail: julian.bertschinger@pharma.ethz.ch

We present a novel method for the directed evolution of polypeptides, which combines *in vitro* compartmentalization and covalent DNA display. A library of linear DNA fragments is co-packaged with an *in vitro* transcription/translation mixture in the compartments of a water-in-oil emulsion. Experimental conditions are adjusted so that, in most cases, one compartment contains one DNA molecule. The DNA fragments encode fusion proteins containing a DNA-methyltransferase (M.Hae III), which can form a covalent bond with a 5-fluorodeoxycytidine base at the extremity of the DNA fragment. The resulting library of DNA–protein fusions is extracted from the emulsion and DNA molecules displaying a protein with desired binding properties are selected from the pool of DNA–protein fusions by affinity panning on target antigens. We applied this methodology in model selection experiments, using specific ligands for the capture of peptides and globular proteins bound to DNA. We observed enrichment factors >1000-fold for selections performed in separate emulsions and up to 150-fold for selections performed using mixtures of DNA molecules. M.Hae III could be fused to small globular proteins (such as calmodulin and fibronectin domains), which are ideally suited for the generation of combinatorial libraries and for the isolation of novel binding specificities.
Keywords: compartmentalization/covalent/DNA display/*in vitro*/selection

Introduction

Protein-based recombinant biopharmaceuticals represent a growing portion of the pharmaceutical market. By the end of 2000, 84 recombinant biopharmaceuticals had gained marketing authorization in the USA and/or the European Union for use in humans (Walsh, 2000). Over the following 3 years, over 60 additional products have been approved and around 500 are undergoing clinical evaluation (Walsh, 2003). Approximately 250 million people have benefited from these 'biotech' medicines and so has the business of biotechnology. The annual global market for biopharmaceuticals is estimated at more than $30 billion, compared with $12 billion just 3 years ago. These numbers clearly indicate the potential benefit of therapeutic proteins. The development of second generation protein drugs with improved therapeutic characteristics such as potency, low immunogenicity, good pharmacokinetics, high stability and expression levels in simple production systems may require the availability of advanced screening and selection methods and improved computational algorithms for the design of tailor-made proteins.

The ability to confer a desired binding specificity to a given protein is of particular interest for pharmaceutical applications, as suitable binding characteristics may improve relevant protein features such as pharmacokinetics, pharmacodynamics and potency and allow a preferential accumulation of the protein at sites of disease.

The engineering of the properties of a protein can be performed by generating a large pool of protein mutants, followed by the identification of those polypeptides which have acquired the desired biological activity, as a consequence of the amino acid mutations accumulated. With recombinant DNA technologies, it is relatively straightforward to generate billions of protein mutants; yet screening all the protein candidates remains a formidable challenge, even in the presence of sophisticated robotic equipment. As a consequence, a considerable part of protein engineering research has focused on the development of efficient selection methodologies, for the rapid identification of proteins with desired characteristics (e.g. binding, catalysis, stability), out of very large pools of candidate proteins.

All selection systems require a link between genotype (a nucleic acid that can be replicated) and phenotype (a functional trait such as binding or catalytic activity). This linkage is usually achieved by physically linking genes to the proteins they encode by a variety of techniques. They can be divided in two classes, which (i) include living organisms in the selection process, as is the case in phage display (Winter *et al*., 1994), yeast surface display (Feldhaus *et al*., 2003) and bacterial surface display (Olsen *et al*., 2000), or (ii) take place fully *in vitro*, such as ribosome display (Mattheakis *et al*., 1994; Schaffitzel *et al*., 1999), mRNA display (Roberts and Szostak, 1997), CIS display (Odegrip *et al*., 2004), DNA display (Tabuchi *et al*., 2001) and selection strategies involving *in vitro* compartmentalization (Tawfik and Griffiths, 1998; Sepp *et al*., 2002; Griffiths and Tawfik, 2003) in combination with DNA display (Doi and Yanagawa, 1999; Yonezawa *et al*., 2003).

The first class of selection methodologies involving live organisms is well established and relatively robust. However, all these methodologies require the transformation of living cells with suitable plasmid DNA, a process which can be laborious and limit library size. A further limiting factor may be represented by the number of cells (or phage particles), which can be used in practical selection procedures.

In order to circumvent the transformation of living cells, research has been focused on the development of the second class of selection strategies taking place fully *in vitro*, of which ribosome display is a prominent example. However, the ternary complex of mRNA, ribosome and displayed polypeptide is stable only under certain conditions (high salt concentrations, low temperatures), thus restricting the experimental conditions in which selections can be performed. In addition, RNA is very

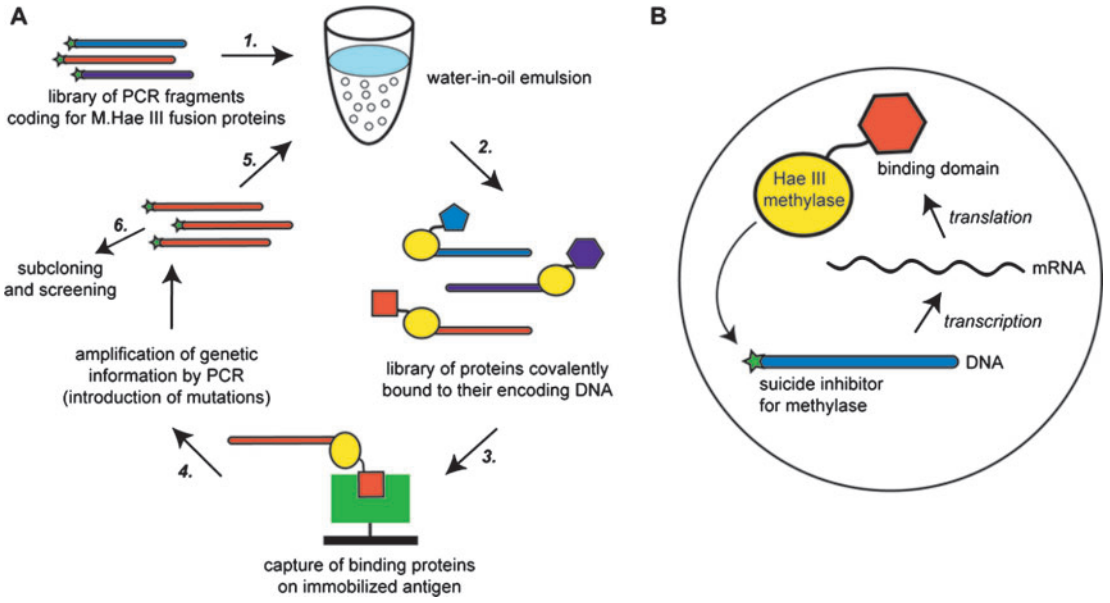

Fig. 1. (**A**) Concept of the selection method. A library of linear DNA molecules is co-packaged with an *in vitro* transcription/translation mix into a water-in-oil emulsion (1). Ideally, one compartment contains one DNA molecule. The DNA molecules each code for a Hae III DNA-methyltransferase fusion protein and contain a mechanism-based inhibitor for the covalent cross-linking of the DNA-methyltransferase fusion proteins [see also (**B**) and Figure 2B]. After *in vitro* expression and formation of the DNA–protein complexes, the water phase is extracted from the emulsion (2) and DNA molecules displaying a protein with desired binding properties are selected from the pool of DNA–protein fusions by affinity panning (3). The genetic information of selected DNA–protein fusions is amplified by PCR (4) and can either be used for a further round of selection (5) or for cloning and characterization of the selected mutants (6). (**B**) Enlarged view of a water compartment of the water-in-oil emulsion. The DNA molecule is transcribed into mRNA, which is translated into a fusion protein consisting of two domains: the N-terminal Hae III DNA-methyltransferase (yellow) and a C-terminal potential binding domain (red). Owing to the catalytic activity of the Hae III DNA-methyltransferase, the fusion proteins form a covalent bond via the modified methylation target sequence 5′-GGFC-3′ (F = FdC = 5-fluorodeoxycytidine) with their encoding DNA molecule, which is present in the same compartment of the water-in-oil emulsion.

sensitive to enzymatic contamination, which may lead to a rapid degradation of the RNA.

We reasoned that an ideal *in vitro* system for the directed evolution of proteins would comprise the following features: (i) rapid generation of genetic diversity (library), (ii) facile and irreversible linkage between phenotype (protein) and genotype (genetic material), (iii) high expression levels of protein mutants in the format used for the selection process, (iv) a carrier of genetic information, which is stable under a wide range of conditions (e.g. high and low salt concentrations, extremes of temperature and pH, denaturing agents) and (v) control of the number of proteins to be linked to the genetic material (valence of display).

In this paper we propose a technology (Bertschinger and Heinis, 2003) in which a library of linear DNA molecules is co-packaged with an *in vitro* transcription/translation mix in the compartments of a water-in-oil emulsion (Tawfik and Griffiths, 1998) (Figure 1A and B). More than a decade ago cell-free translation was achieved in reversed micelles of surfactants in organic solvents (Nametkin *et al*., 1992). A few years later, Tawfik and Griffiths (1998) showed in their pioneering work that the composition of the organic phase of emulsions could be adjusted to achieve stable, cell-like compartments for *in vitro* compartmentalization. Experimental conditions are chosen so that, in most cases, one compartment contains one DNA molecule (Figure 2A). A covalent linkage between DNA and the encoded polypeptide is achieved by using fusion proteins which contain the Hae III DNA-methyltransferase domain of *Haemophilus aegypticus*, an enzyme which is able to form a covalent bond with DNA fragments containing the sequence 5′-GGFC-3′ (F = 5-fluoro-2′-deoxycytidine) (Chen *et al*., 1991) (Figure 2B).

M.Hae III fusion proteins expressed in each compartment of the water-in-oil emulsion are extracted from the emulsion. The DNA molecules which display a protein with the desired binding specificity are selected from the pool of DNA–protein fusions by affinity panning. The genetic information of the selected DNA–protein fusions is amplified by polymerase chain reaction (PCR) and can be used either for a further round of selection or for cloning and characterization of the selected mutants (Figure 1A).

Our experiments show that the ability of Hae III DNA-methyltransferase to form a covalent complex with DNA molecules containing the suicide inhibitor sequence 5′-GGFC-3′ can be exploited to create a stable linkage between genotype and phenotype *in vitro*. In model selection experiments, DNA molecules were enriched on the basis of the protein they encode from an excess of DNA molecules coding for a protein with irrelevant binding specificity. Hae III DNA-methyltransferase can be fused to small globular proteins (such as calmodulin or fibronectin domains) which are devoid of disulfide bridges and lend themselves to the generation of combinatorial libraries for the isolation of specific binding proteins (Smith, 1998).

Materials and methods

Preparation of water-in-oil emulsions and determination of size distributions of the water compartments

Water-in-oil emulsions were essentially prepared as described by Tawfik and Griffiths (1998). A 10-fold concentrated solution of oil phase was prepared by mixing 50% (v/v) mineral oil (Sigma-Aldrich) with 45% (v/v) Span 80 (Fluka, Buchs, Switzerland) and 5% (v/v) Tween 80 (Fluka). For the

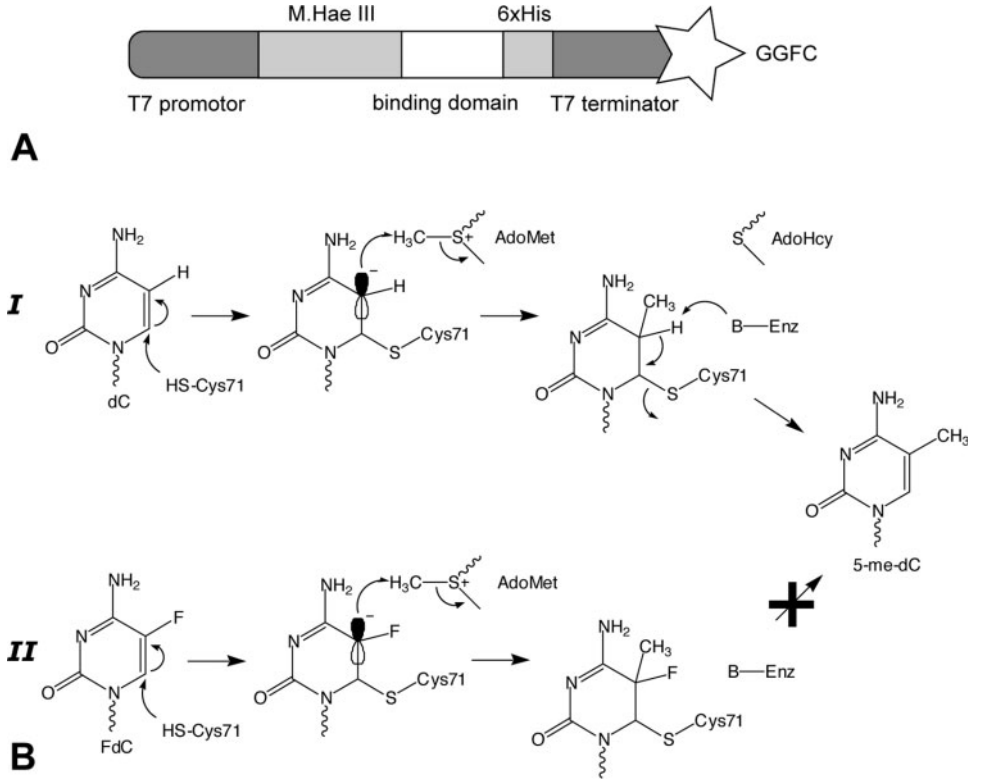

Fig. 2. (**A**) Schematic representation of DNA template used for the *in vitro* expression of M.Hae III fusion proteins. The sequences coding for Hae III DNA-methyltransferase and the potential binding domain are flanked by T7 regulatory sequences (T7 promoter and T7 terminator) for efficient expression. The modified base 5-FdC is introduced into the DNA by PCR at the downstream end of the DNA T7 terminator sequence. (**B**) Chemical reaction mechanisms that lead either to the methylation of deoxycytidine (dC) (I) or to the covalent linkage of M.Hae III with the modified methylation sequence 5′-GGFC-3′ (F = FdC = 5-fluorodeoxycytidine) (II). *I:* Methylation is initiated by nucleophile attack of an active site thiol group (Cys71) at C6 of cytosine. *S*-Adenosylmethionine (AdoMet) acts as the methyl donor. The methylated dihydrocytosine intermediate undergoes beta-elimination to generate a 5-methylcytosine and an active enzyme. *II:* If the nucleotide 5-fluorodeoxycytidine is inserted in the methylation sequence (5′-GGFC-3′), the dihydrocytosine beta-elimination is blocked. The enzyme remains covalently bound to the DNA substrate.

preparation of the emulsions, this concentrated stock solution was diluted 10-fold in mineral oil and 950 µl were used for the encapsulation of 50 µl of water phase. The mixing of the water and the organic phases was done by the incremental addition of 5× 10 µl of ice-cooled water phase to ice-cooled oil phase in a glass vial (Forma Vitrum, Switzerland, 40 × 18.75 mm). Stirring was done at 2200 r.p.m. using a magnetic stirrer (MR 1000, Heidolph, Germany) and a stirring bar (3 mm) with a pivot ring. After addition of water phase, stirring was continued for further 5 min to obtain the desired size distribution of the water compartments.

The size distributions of the water compartments were measured by laser diffraction using a Mastersizer X (Malvern Instruments, UK). A 30 µl volume of an emulsion was added to 10 ml of mineral oil and mixed in the presentation cell of the Mastersizer (30–40% stirring power). Particle size distributions were calculated with the instrument's built-in Fraunhofer optical model (Figure 3A) or Mie optical model (Figure 3B) using refractive indices of 1.470 for mineral oil and 1.330 for the aqueous phase.

Formation of covalent DNA–protein adducts

A 268 bp DNA fragment containing the suicide inhibitor sequence 5′-GGFC-3′ for M.Hae III was used for the cross-linking experiments. The modified nucleotide 5-fluorodeoxycytidine was introduced by PCR amplification (T3 Thermocycler, Biometra) of the DNA supplied in vial 3 of an RTS *Escherichia coli* Linear Template Generation Set (Roche Applied Sciences) using the synthetic oligonucleotides Hae sub fo (5′-C GTC AT**G GFC** TAT GCG GGC GAC CAC ACC CGT CCT GTG GAT-3′, modified substrate sequence in bold letters) and Hae sub ba 1 (5′-TAT TGC TAT GTA CCT AGC GGG GGG GGT TCT CAT CAT CAT-3′) (both oligonucleotides from Microsynth, Switzerland). A 2 nM amount of DNA was incubated with 38 nM M.Hae III (New England Biolabs, USA) and 80 µM *S*-adenosylmethionine (New England Biolabs) in DNA-methyltransferase reaction buffer (50 mM NaCl, 50 mM Tris–HCl (pH 8.5), 10 mM dithiothreitol; New England Biolabs). Cross-linking reactions were stopped by heat inactivation (15 min at 70°C). Samples were loaded on a denaturing 10% TBE urea polyacrylamide gel (Invitrogen) and DNA was visualized with SYBR green II (Molecular Probes, USA). A 10 bp DNA Ladder (Invitrogen) was used as a DNA marker.

Preparation of templates for in vitro transcription and translation

The gene coding for Hae III DNA-methyltransferase was ordered at ATCC (strain 11116). The DNA templates for the expression of the various M.Hae III constructs were assembled following the instructions of the RTS *E.coli* Linear Template Generation Set (Roche Applied Sciences) with the exception that the oligonucleotides Ext ba 2 (5′-GAT GCC GGC CAC GAT GCG TCC GGC GTA GAG-3′, Qiagen) and Ext 2 fo (5′-GCT AAT TA**G GFC** ACC ACA CCC GTC CT-3′, modified substrate sequence in bold letters; Microsynth, Switzerland) were used for the final PCR assembly

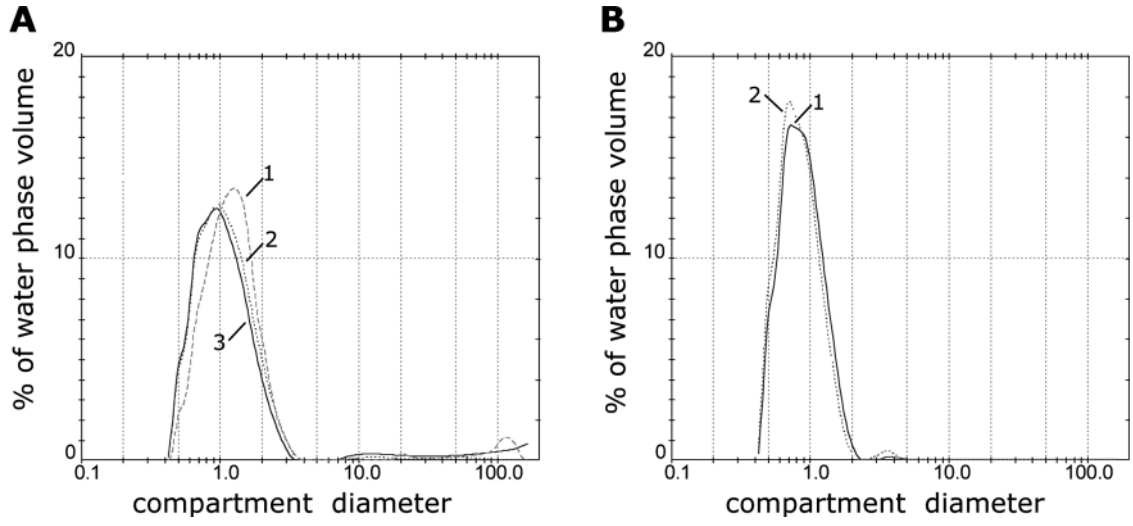

Fig. 3. (**A**) Reproducibility of emulsion preparation: three emulsions were independently prepared as described above (curves 1, 2 and 3). The size distributions were very similar. (**B**) Stability of water-in-oil emulsions: an emulsion was prepared and size distributions of water compartments were measured after storage for 0 h (1) and 96 h (2) at room temperature, respectively.

instead of the primers supplied with the Linear Template Generation Set. DNA was purified using a QIAquick Gel extraction kit or QIAquick PCR purification kit (Qiagen). If DNA was used directly for *in vitro* transcription/translation reactions, DNA samples never were purified over an agarose gel because agarose inhibits *in vitro* protein expression (Manual RTS *E.coli* HY kit, Roche Applied Sciences).

Further details concerning the assembly of the different M.Hae III fusions are as follows. M.Hae III-Flag tag: the amino acid residues GGSG-DYKDDDK followed by two stop codons were appended to the DNA-methyltransferase gene in order to avoid the expression of the 6 × His tag, which is included in the Linear Template Generation Set. M.Hae III-EDB/CaM: the sequences coding for calmodulin (CaM) and the extra-domain B of fibronectin (EDB) were available in our laboratory. The amino acid residues GGSGAS were used as linker for the C-terminal fusion of EDB/CaM to M.Hae III. GST-M.Hae III: the M.Hae III gene was cloned as a C-terminal fusion to glutathione-*S*-transferase (GST) in the vector pGEX-4T-2 (Amersham Biosciences) using the restriction sites *Bam*HI and *Not*I (both enzymes from Promega). Both genes were amplified together from the resulting vector and the PCR product was used for PCR assembly using the Linear Template Generation Set.

In vitro transcription/translation and western blot

All *in vitro* transcription/translation reactions were done using the RTS 100 *E.coli* HY kit (Roche Applied Sciences). In order to allow the formation of covalent DNA–protein adducts, 80 μM *S*-adenoslymethionine was added to the transcription/translation mix. For western blot analysis, 2 μl of the transcription/translation mixtures (non-emulsified expression) were loaded on a 10% Bis-Tris SDS–polyacrylamide gel [Invitrogen, protein marker RPN 800 (Amersham Biosciences)] followed by blotting on a PVDF membrane (0.45 μm pores, Immobilon Transfer Membrane, Millipore), which was previously activated by soaking for 15 s in 100% methanol followed by 2 min in water. After blotting, the PVDF membrane was quickly rinsed with water, soaked for 2 × 15 s in methanol and then air-dried on a filter-paper. M.Hae III fusion proteins bearing a 6 × His tag were detected using an anti-His-HRP immunoconjugate (Sigma-Aldrich). M.Hae III-Flag tag protein was detected using anti-Flag tag (M2) as primary antibody (Sigma-Aldrich) and anti-mouse-HRP immunoconjugate as secondary antibody (Sigma-Aldrich). The PVDF membrane was incubated with the antibodies for 45 min at room temperature in 3% milk PBS (shaking at 65 r.p.m.). Subsequently, the membrane was washed four times with PBS. HRP activity was detected with the ECL kit (Amersham Biosciences) as indicated by the manufacturer.

Model selection experiments

In vitro transcription/translation mix (TS/TL-mix) was prepared on ice according to the instructions of the manufacturer of the kit (RTS 100 *E.coli* HY kit, Roche Applied Sciences). A 40 μl volume of the TS/TL-mix, 5 μl of *S*-adenosylmethionine (stock, 800 μM; final concentration, 80 μM), DNA and 5 μl of water were mixed in order to obtain a final volume of 50 μl. DNA was added just before preparing the emulsion. A 50 μl volume of ice-cooled water phase (*in vitro* TS/TL-mix containing the DNA template) was added to 950 μl of oil phase (ice-cold) in a stepwise manner (5 × 10 μl during 2 min) whilst stirring at 2200 r.p.m. in a glass vial. After stirring, the emulsion was transferred to a 1.5 ml microcentrifuge tube.

Proteins were expressed in emulsion by incubation for 150 min in a heat block at 30°C. After protein expression and formation of the covalent DNA–protein adducts, emulsions were broken by centrifugation for 10 min at 7000 r.p.m. in a table-top centrifuge (Eppendorf). Water droplets formed a pellet and the organic supernatant was removed. A 150 μl volume of room temperature breaking buffer was added [PBS (or TBS with 1 mM $CaCl_2$ for selections with calmodulin (TBSC)], pH 7.4, 5 μM biotinylated ds DNA fragments [5′-biotin-GGA GCT TCT GCA TTC TGT GTG CTG-3′ (Qiagen)], 1 μM competing ds DNA fragments [5′-ATC TAA **GGC** **C**AA TGT ACT AGA C**GG** **CC**A TTC CAG ATG CA**G** **GCC** AAG CGT ACA TAC **GGC** **C**TA GCT AAA TCA A**GG** **CC**G TAT CGT-3′, substrate recognition sites of M.Hae III in bold letters, (Qiagen)] followed by 1 ml of room temperature diethyl ether (water saturated). The tube

was vortexed at full speed for 2 × 10 s. After separation of the water phase and the organic phase, the water phase was removed from the bottom of the tube and transferred to a 24-well plate (NUNCLON Surface, Nunc). The extracted water phase was air-dried at room temperature for 10 min until residual Et_2O was completely evaporated.

Volumes of 25–50 µl of magnetic streptavidin-coated Dynabeads (Dynal, Norway) were incubated for 15 min with the capture agent at room temperature {biotinylated anti-penta-His antibody (4 µl or 800 ng/50 µl beads) (Qiagen), biotinylated anti-Flag antibody (M2, 2 µl or 2 µg/50 µl beads) (Sigma-Aldrich), biotinylated anti-EDB antibody (5 µl/50 µl beads, 200 nM) [L19 small immunoprotein (Li *et al*., 1997), which was kindly provided by P.Alessi, Institute of Pharmaceutical Sciences, ETH Zurich, Switzerland], biotinylated calmodulin binding peptide [biotin-CAAARWKKAFIAVSAANRFKKIS (Montigiani *et al*., 1996), which was kindly provided by L.Lozzi, Dipartimento di Biologia Molecolare, Universita' di Siena, Italy] (4 µl/50 µl beads, 800 nM)}. The beads were washed once with PBS 0.1% Tween 20 (PBST) or TBSC 0.1% Tween 20 (TBSCT) (if CaM-binding peptide was used as capture agent) and blocked with 5 µM biotinylated ds DNA fragments [5′-biotin-GGA GCT TCT GCA TTC TGT GTG CTG-3′ (Qiagen)] for 15 min at room temperature. The beads were then incubated with the extracted water phase for 45 min at room temperature. The samples were shaken gently from time to time. The magnetic beads were washed six times with 100 µl of PBST or TBSCT, followed by one wash with 100 µl of PBS or TBSC using a magnetic separator (Dynal, Norway). The beads were resuspended in 100 µl of water and used directly for PCR amplification of the DNA molecules remaining on the beads using the primers Ampl ba (5′-CCC GCG AAA TTA ATA CGA CTC A-3′, Qiagen) and Ampl fo (5′-AAA ACC CCT CAA GAC CCG TT-3′, Qiagen). As an exception, PCRs for Figure 6 were done using the primers M.Hae Nco ba (5′-GGA GAT ATA ACC ATG GGC AAT TTA ATT AGT CTT TTT TCA GGT-3′, Qiagen) and M.Hae Xho His fo (5′-TAT AGC TCG CTC GAG ATT ACC TTT ACA AAT TTC CAA TGC AGA-3′, Qiagen). Smart Ladder (Eurogentec) was used as a DNA marker on the agarose gel (Figure 6).

Real-time PCR

An ABI PRISM 7000 instrument (Applied Biosystems), ABI PRISM 96-well Optical Reaction Plates with MiniAmp Optical Caps (Applied Biosystems) and TaqMan Universal PCR Master Mix (Applied Biosystems) were used for all real-time PCR measurements. All reactions were done in triplicate for each sample. Primers and TaqMan probes were designed with Primer Express 1.5 software (Applied Biosystems). Primers were used at a concentration of 200 nM and the TaqMan probe at 100 nM. DNA samples were diluted in such a way that the expected number of molecules detected by real-time PCR was between 10^2 and 10^6.

Results

Preparation of water-in-oil emulsions and determination of size distribution of the water compartments

A basic requirement for the proposed methodology is the ability to produce a water-in-oil emulsion with individual compartments, whose size and stability (no fusion with other compartments) can be controlled. The water-in-oil emulsions were prepared by stirring a mixture of oil phase and water phase. The latter contained the DNA template and an *in vitro* transcription/translation mix. The time and the velocity of stirring are critical parameters for the shape of the size distribution of the water compartments of the water-in-oil emulsion. Water-in-oil emulsions were prepared as described in the literature (Tawfik and Griffiths, 1998) with minor modifications and the size distribution of the water compartments of the emulsions was determined by laser diffraction. Three emulsions were prepared independently and the corresponding size distributions were measured (Figure 3A). The average diameter of the water compartments was about 1 µm with 67.5% (±3.5% s.d., $n = 3$) of the water volume being in compartments with diameters ranging from 0.71 to 1.84 µm (area under the curve).

In order to determine the stability of the water-in-oil emulsions, we checked whether the size distribution of the water compartments changed over time. One emulsion was prepared and the size distribution of the water compartments was determined immediately using only a small amount of the emulsion. The remaining part of the emulsion was left for 96 h at room temperature until the second measurement was performed (Figure 3B). The profiles of the two curves obtained at the different time points did not differ significantly from each other, indicating a remarkable stability of the water-in-oil emulsions.

Formation of covalent protein–DNA adducts

We showed that M.Hae III can be linked covalently to DNA fragments which contain the modified substrate sequence for the DNA-methyltransferase (5′-GGFC-3′) by incubating DNA (2 nM), *S*-adenosylmethionine (SAM, 80 µM) and M.Hae III (38 nM) for different time periods (Chen *et al*., 1991). The concentrations of the reactants were chosen to be in the low nanomolar range in order to mimic the conditions for cross-linking DNA with protein, assuming that one DNA molecule with about 15 molecules of M.Hae III are present in one compartment of a water-in-oil emulsion after transcription and translation. Reactions were stopped by heat inactivation of the DNA-methyltransferase (15 min at 70°C) and the outcome of the experiment was analysed on a denaturing TBE gel containing 8 M urea (Figure 4). The heating step and the

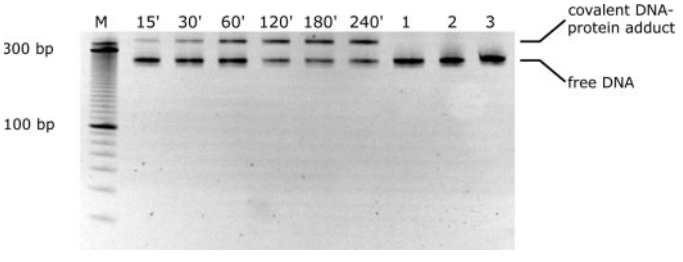

Fig. 4. Covalent coupling of M.Hae III with a 268 bp DNA fragment containing the modified methylation site 5′-GGFC-3′. DNA (2 nM), M.Hae III (38 nM) and *S*-adenosylmethionine were incubated at 37°C for different time periods (given in minutes above the gel). M.Hae III was inactivated by heating (70°C, 15 min) and the samples were loaded on a denaturing polyacrylamide gel (10%, 8 M urea) and DNA was stained with SYBR green II. DNA molecules covalently bound to M.Hae III are shifted upwards as compared with free DNA (indicated in the figure). Three controls are shown on the right: 1, without *S*-adenosylmethionine; 2, without M.Hae III; 3, only 268 bp DNA fragment; M, 10 bp marker.

denaturing conditions during gel electrophoresis should allow the disruption of all non-covalent protein–DNA complexes. After a reaction time of 120 min, about 50% of the DNA molecules were covalently associated with one protein of M.Hae III DNA-methyltransferase. The yield of DNA–protein adducts did not increase if the reaction was allowed to proceed for up to 240 min. If DNA containing the suicide inhibitor and Hae III DNA-methyltransferase were incubated without the cofactor SAM, no covalent protein–DNA adducts were detected. This result is in accordance with the reaction mechanism described in the literature (Figure 2B) (Chen *et al*., 1991).

In vitro expression of M.Hae III–polypeptide fusions

A further important requirement for a selection methodology is the possibility of expressing a variety of different polypeptides in the format needed for the selection process. Therefore, we examined different M.Hae III fusion proteins for their expression levels *in vitro*. Five DNA templates were assembled by PCR coding for M.Hae III-6 × His tag, M.Hae III-Flag tag, M.Hae III-CaM-6 × His (calmodulin), M.Hae III-EDB-6 × His (extra-domain B of fibronectin) and GST-M.Hae III-6 × His (glutathione-*S*-transferase of *Schistosoma japonicum*). All polypeptides except GST were C-terminal fusions to M.Hae III.

Each M.Hae III fusion protein was expressed using 200–250 ng of DNA template in a volume of 25 µl of *in vitro* transcription/translation mixture and, after incubation at 30°C for 3.5 h, the expression levels of the fusion proteins were analysed by western blotting. Proteins were detected using either anti-His tag–horseradish peroxidase (HRP) immunoconjugate (for M.Hae III-6 × His tag, -CaM-6 × His tag, -EDB-6 × His tag and GST-M.Hae III-6 × His tag) or anti-Flag antibody (M2) in combination with a secondary anti-mouse-HRP detection reagent (for M.Hae III-Flag tag) (Figure 5). All constructs except GST-M.Hae III-6 × His were well expressed and appeared to have the correct size, without any detectable proteolytic fragments. The fusion proteins also were catalytically active, as shown by the ability of the fusion proteins to protect DNA fragments from digestion with the restriction nuclease *Not*I (data not shown). The expression yield for M.Hae III-6 × His tag was estimated from signal intensities on western blots to be about 300 ng/µl *in vitro* transcription/translation mixture (data not shown). Expression of M.Hae III fusion proteins was markedly lower (~20%) compared with expression yields obtained in bulk solution (data not shown).

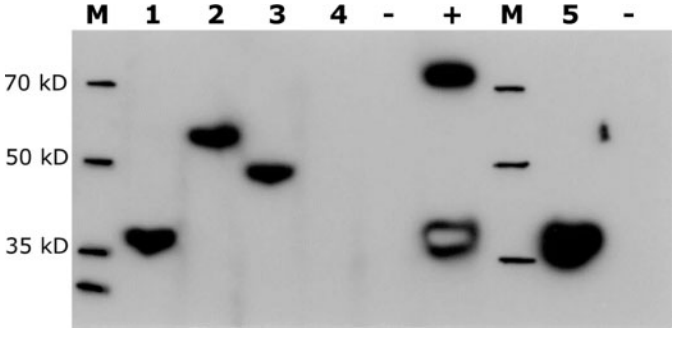

Fig. 5. *In vitro* expression of M.Hae III fusion proteins. Linear DNA templates coding for different M.Hae III fusion proteins were incubated in an *in vitro* transcription/translation mix. A 2 µl volume was taken for western blot analysis [detection: anti-His-HRP or anti-Flag (M2)/anti-mouse-HRP]. M, marker; 1, M.Hae III-6 × His tag (total molecular weight, 38 kDa); 2, M.Hae-calmodulin-6 × His (56 kDa); 3, M.Hae III-ED-B-6 × His (extra-domain B of fibronectin) (49 kDa); 4, GST-M.Hae III-6 × His tag (*S.japonicum*) (64 kDa); 5, M.Hae III-Flag tag (38 kDa); +, 1.5 µg recombinantly expressed M.Hae III-6 × His tag (38 kDa); –, transcription/translation mix. All peptides/proteins except GST were C-terminal fusions to M.Hae III.

Model affinity selections with DNA templates coding for M.Hae III-6 × His or -Flag fusion proteins

In order to test the practical feasibility of covalent DNA display methodology, we performed experiments using two peptides (6 × His and Flag tag) as M.Hae III fusion and two specific monoclonal antibodies for affinity capture. Amounts of 100 ng of DNA molecules coding either for M.Hae III-6 × His tag or M.Hae III-Flag tag protein were emulsified separately. After protein expression at 30°C for 2.5 h, the two emulsions were centrifuged for 10 min at 10 000 r.p.m.. The oil supernatants were removed and the water phases, containing the protein–DNA adducts, were extracted from the pellets formed by the still intact water compartments at the bottom of the tubes by addition of diethyl ether. Subsequently, the extracted water phases were dried to remove residual traces of diethyl ether and divided into two halves each. The first two halves of the water phases containing DNA-M.Hae III-6 × His or DNA-M.Hae III-Flag adducts were used for affinity pannings with biotinylated anti-penta-His and the second two halves were taken for pannings with biotinylated anti-Flag antibody immobilized on magnetic streptavidin beads. DNA molecules still bound to the beads after washing were amplified by PCR (amplicon length = 1020 bp) without prior elution from the magnetic beads. The outcome of the selections was analysed on an agarose gel (Figure 6). As expected, only DNA coding for M.Hae III-6 × His was selected if the affinity capture was performed with anti-penta-His antibody, whereas DNA coding

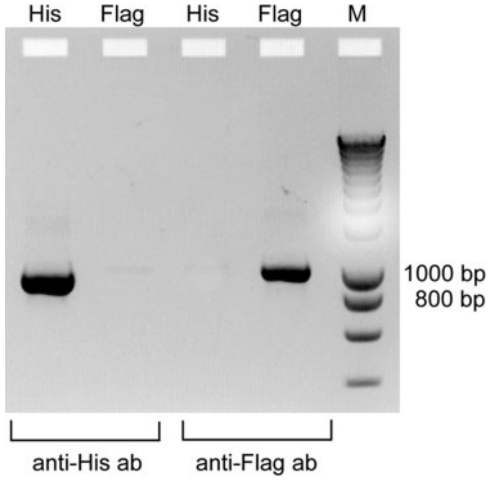

Fig. 6. Model selections: 100 ng linear DNA templates (6.6×10^{10} molecules) coding for either M.Hae III-6 × His tag or M.Hae III-Flag tag were emulsified with 50 µl of *in vitro* transcription/translation mixture. Covalent protein–DNA adducts were extracted from the emulsions after expression for 2 h at 30°C. Affinity panning experiments were done with magnetic streptavidin beads coated either with biotinylated anti-His or anti-Flag (M2) antibody. In order to amplify the selected DNA molecules, PCR was performed directly on the beads after washing. PCR products (1020 bp long) were loaded on an agarose gel. His = DNA coding for M.Hae III-6 × His tag used for selection; Flag = DNA coding for M.Hae III-Flag tag used for selection. Affinity panning with anti-His antibody or anti-Flag antibody as indicated below the agarose gel. M, DNA marker.

Table I. Affinity selections with different M.Hae III fusion constructs

M.Hae III construct	Input	Capture agent	Output[a] (±s.d.[b])	Ratio[c]
M.Hae III-Flag tag	6×10^{10}	anti-Flag ab	$2.3 \pm 0.2 \times 10^{6}$	
M.Hae III-6 × His	6×10^{10}		$7.6 \pm 2.0 \times 10^{2}$	3026
M.Hae III-Flag tag	6×10^{10}	anti-Flag ab	$4.0 \pm 0.5 \times 10^{5}$	
M.Hae III-6 × His	6×10^{10}		$5.0 \pm 0.8 \times 10^{2}$	800
M.Hae III-Flag tag	6×10^{10}	anti-Flag ab	$1.7 \pm 0.1 \times 10^{6}$	
M.Hae III-6 × His	6×10^{10}		$5.5 \pm 1.4 \times 10^{2}$	3091
M.Hae III-Flag tag	6×10^{10}	anti-Flag ab	$1.9 \pm 0.2 \times 10^{6}$	
M.Hae III-6 × His	6×10^{10}		$5.2 \pm 1.6 \times 10^{2}$	3654
M.Hae III-6 × His	3×10^{10}	anti-His ab	$4.4 \pm 0.3 \times 10^{5}$	
	3×10^{10}	anti-Flag ab	$7.9 \pm 3.6 \times 10^{2}$	557
M.Hae III-6 × His	1×10^{10}	anti-His ab	$2.0 \pm 0.1 \times 10^{6}$	
	1×10^{10}	anti-Flag ab	$2.9 \pm 0.9 \times 10^{2}$	6897
M.Hae III-EDB	5×10^{10}	anti-EDB ab (L19)	$1.3 \pm 0.1 \times 10^{6}$	
	5×10^{10}	anti-Flag ab	$4.0 \pm 0.2 \times 10^{2}$	3250
M.Hae III-CaM	5×10^{10}	CaM-pep	$2.9 \pm 0.2 \times 10^{5}$	
	5×10^{10}	anti-Flag ab	$1.7 \pm 0.6 \times 10^{2}$	1076
M.Hae III-CaM	5×10^{10}	CaM-pep	$7.8 \pm 1.1 \times 10^{5}$	
	5×10^{10}	anti-Flag	$6.9 \pm 1.4 \times 10^{2}$	1130

[a]Output values were measured by real-time PCR using 0.1% of the beads after washing.
[b]Standard deviation: all measurements were performed in triplicate.
[c]Ratio of the two corresponding output values.

for M.Hae III-Flag tag could not be detected. However, if anti-Flag antibody was used for affinity selection, only DNA coding for M.Hae III-Flag was recovered and detected on the agarose gel after PCR. Therefore, we concluded that it is possible to select DNA molecules via the binding specificity of the encoded protein.

Quantification of selection efficiency

For the measurement of the efficiency with which DNA–protein adducts were selected by virtue of their binding affinity to the immobilized antibody on the magnetic beads, model selection experiments were performed as described above using additional M.Hae III fusion constructs. The number of DNA molecules recovered was determined by real-time PCR using a TaqMan probe specific for the DNA-methyltransferase gene. In a first set of experiments (Table I, upper half), two emulsions were prepared, one of which contained DNA coding for M.Hae III-6 × His and the other DNA encoding M.Hae III-Flag. After protein expression and extraction of the water phases, affinity selections using magnetic streptavidin beads coated with biotinylated anti-Flag antibody were performed. In Table I, the number of DNA molecules used as input for the selections and the corresponding outputs are given. The data show that DNA-M.Hae III-Flag adducts were selected ~1000 times more efficiently than an irrelevant DNA–protein adduct (in this case M.Hae III-6 × His).

In a second set of experiments, various M.Hae III constructs were separately expressed in emulsion and, after extraction from the emulsion, affinity selections were performed using magnetic streptavidin beads coated with specific or irrelevant antibody. If the streptavidin beads were coated with specific antibody/peptide, DNA molecules were recovered much more efficiently than compared to the control selection, in which an irrelevant antibody was immobilized on the beads (Table I, lower half).

Model selections with DNA mixtures

For practical applications, the covalent DNA display methodology will be preferentially used for the selection of globular proteins of desired binding specificities. In the ideal case, such proteins are devoid of disulfide bonds, are stably folded and are well expressed in bacteria. Calmodulin (Neri *et al*., 1995; Montigiani *et al*., 1996) and a number of fibronectin domains (Zardi *et al*., 1987; Xu *et al*., 2002) fulfil these requirements. Mixtures of DNA templates coding either for M.Hae III-CaM or M.Hae III-EDB were prepared in such a way that the ratio of M.Hae III-EDB/-CaM DNA molecules was ~1000 and that 10^{9}–10^{10} DNA molecules were used per selection experiment. Biotinylated CaM-binding peptide was used to coat the magnetic beads in order to enrich DNA coding for M.Hae III-CaM from these DNA mixtures, whereas biotinylated anti-Flag antibody served as a negative control because neither of the two proteins contained a Flag-tag and, therefore, the ratio of M.Hae III-EDB/-CaM DNA molecules should not change as compared with the input ratio. DNA molecules recovered from the selection process were amplified by PCR. Selection experiments were analysed by real-time PCR using TaqMan probes specific for either M.Hae III-EDB or -CaM DNA. The ratio of M.Hae III-EDB/-CaM DNA molecules was determined before selection (input) and after selection with CaM-binding peptide or anti-Flag antibody (Table II). In the last column of Table II, enrichment factors were calculated by dividing the value of the input ratio by the value of the output ratio (using CaM-binding peptide as capture agent). The DNA mixtures, obtained from PCR amplification of recovered DNA molecules on the magnetic beads, were diluted $1{:}10^{4}$–10^{5} for real-time PCR analysis.

Using 10^{10} DNA molecules as input for the selection procedure, enrichment factors ranged between 42 and 94 (Table II, selection experiments 1–7). If only 10^{9} DNA molecules were used, enrichment factors increased to 81–153 (Table II, selection experiments 8–10). In some control selections with the irrelevant anti-Flag antibody, DNA coding for M.Hae III-CaM was not detectable by real-time PCR, having been diluted out during the washing steps.

Discussion

We propose an *in vitro* selection system in which phenotype (protein) and genotype (DNA) are physically linked by a

Table II. Model selections with mixtures of M.Hae III–EDB and –CaM DNA molecules

Exp.	No. of DNA molecules[a]	Input EDB	Input CaM	Input ratio EBD/CaM[b]	Output EDB anti-Flag	Output CaM anti-Flag	Ratio anti-Flag[c]	Output EDB CaM-pep	Output CaM CaM-pep	Ratio CaM-pep[d]	Enrichment factor[e]
1	10^{10}	$9.7 \pm 0.7 \times 10^{5}$	$2.1 \pm 0.2 \times 10^{2}$	4619	$9.7 \pm 1.1 \times 10^{6}$	n.d.[f]	n.d.	$3.6 \pm 0.2 \times 10^{6}$	$7.4 \pm 1.6 \times 10^{4}$	49	94
2	10^{10}	$1.2 \pm 0.2 \times 10^{6}$	$2.7 \pm 0.4 \times 10^{2}$	4444	$3.4 \pm 0.3 \times 10^{6}$	n.d.	n.d.	$2.3 \pm 0.1 \times 10^{6}$	$2.6 \pm 0.3 \times 10^{4}$	88	50
3	10^{10}	$1.0 \pm 0.2 \times 10^{6}$	$2.4 \pm 0.2 \times 10^{2}$	4167	$4.9 \pm 0.8 \times 10^{5}$	$5.0 \pm 0.2 \times 10^{2}$	980	$5.5 \pm 0.8 \times 10^{5}$	$5.5 \pm 0.5 \times 10^{3}$	100	42
4	10^{10}	$2.5 \pm 0.3 \times 10^{6}$	$1.5 \pm 0.3 \times 10^{3}$	1667	$7.2 \pm 0.4 \times 10^{5}$	$1.9 \pm 0.6 \times 10^{2}$	3789	$3.5 \pm 0.2 \times 10^{5}$	$9.1 \pm 1.5 \times 10^{3}$	38	43
5	10^{10}	$2.3 \pm 0.3 \times 10^{6}$	$1.3 \pm 0.4 \times 10^{3}$	1769	$1.7 \pm 0.1 \times 10^{6}$	$6.7 \pm 2.1 \times 10^{2}$	2537	$7.0 \pm 0.1 \times 10^{5}$	$1.8 \pm 0.5 \times 10^{4}$	39	45
6	10^{10}	$1.5 \pm 0.2 \times 10^{6}$	$6.8 \pm 1.4 \times 10^{2}$	2206	$6.1 \pm 1.2 \times 10^{5}$	$7.4 \pm 0.9 \times 10^{2}$	824	$8.4 \pm 0.4 \times 10^{5}$	$1.9 \pm 0.2 \times 10^{4}$	44	50
7	10^{10}	$1.5 \pm 0.2 \times 10^{6}$	$6.8 \pm 1.4 \times 10^{2}$	2206	$4.5 \pm 0.9 \times 10^{5}$	n.d.	n.d.	$1.3 \pm 0.5 \times 10^{6}$	$2.8 \pm 0.5 \times 10^{4}$	46	48
8	10^{9}	$8.4 \pm 0.6 \times 10^{5}$	$2.0 \pm 0.1 \times 10^{2}$	4200	$1.5 \pm 0.1 \times 10^{6}$	n.d.	n.d.	$1.4 \pm 0.2 \times 10^{6}$	$5.1 \pm 0.7 \times 10^{4}$	27	153
9	10^{9}	$9.5 \pm 1.6 \times 10^{5}$	$5.1 \pm 1.1 \times 10^{2}$	1863	$7.9 \pm 1.1 \times 10^{4}$	n.d.	n.d.	$1.5 \pm 0.1 \times 10^{5}$	$6.5 \pm 0.1 \times 10^{3}$	23	81
10	10^{9}	$1.3 \pm 0.4 \times 10^{6}$	$4.2 \pm 0.3 \times 10^{2}$	3095	$6.6 \pm 0.9 \times 10^{4}$	n.d.	n.d.	$2.4 \pm 0.3 \times 10^{5}$	$9.2 \pm 1.9 \times 10^{3}$	26	119

[a]Total number of DNA molecules used for selections.
[b]Ratio of M.Hae III-EDB/M.Hae III-CaM DNA molecules used as input; PCR with diluted input mixtures decreased ratio by a factor of 2–3.
[c]Ratio of M.Hae III-EDB/M.Hae III-CaM DNA molecules after selection with anti-Flag antibody as capture reagent.
[d]Ratio of M.Hae III-EDB/M.Hae III-CaM DNA molecules after selection with CaM-binding peptide as capture reagent.
[e]Enrichment = [input ratio]/[output ratio (obtained with CaM-binding peptide)].
[f]n.d.: no DNA molecules were detectable in these samples.

covalent bond. This physical linkage between a protein and its encoding DNA molecule was created by exploiting the ability of Hae III DNA-methyltransferase of *H.aegypticus* to form a covalent bond with DNA fragments containing the sequence 5′-GGFC-3′ (F = 5-fluorodeoxycytidine) (Chen *et al*., 1991). DNA molecules coding for M.Hae III-CaM proteins were successfully and reproducibly selected in model selection experiments from a more than 1000-fold excess of DNA molecules coding for M.Hae III-EDB.

As shown in Table I, the efficiency of DNA recovery from affinity selections varied by more than three orders of magnitude, depending on the binding specificity of the polypeptide used as capture agent. These results imply that DNA molecules coding for a protein with appropriate binding characteristics could be enriched from a mixture of different DNA molecules by a factor of ~1000. However, enrichment factors obtained from model selection experiments starting from mixtures of two different DNA templates were lower than expected by a factor of 10–20 (Table II). These findings may be explained by different reasons.

First, more than one DNA molecule might be present in one water compartment of the water-in-oil emulsion, which would lead to the incorrect assignment of genotype and phenotype and, therefore, to the selection of DNA molecules coding for non-binding proteins. This could explain why enrichment factors increased by almost a factor of three if 10^{9} instead of 10^{10} DNA molecules were used as input for the selection experiments. With the mean diameter of a water compartment of the emulsion being ~1 μm (Figure 3) and given that 50 μl of water phase are encapsulated in 950 μl of oil phase, there should be roughly 10^{11} distinct water compartments in the resulting emulsion. However, this calculation contradicts the results obtained from selections with 10^{10} and 10^{9} DNA molecules. If 10^{11} water compartments were formed in 1 ml of emulsion and 10^{10} DNA molecules were used as input, there would be only 0.1 DNA molecule on average per water compartment. Therefore, one would not expect the enrichment factors to improve if the starting amount of DNA molecules is lowered by a factor of 10. However, enrichment factors improved by decreasing the amount of DNA molecules in the selection and, therefore, the size distributions might be shifted to larger mean diameters (Figure 3), as described previously (Tawfik and Griffiths, 1998). It should be noted that emulsions had to be diluted in mineral oil before size distributions were measured (see Materials and methods), which could influence the compartment size.

Second, fusion of compartments followed by fission taking place during protein expression and formation of DNA–protein adducts could lead to the exchange of the contents between different water compartments without changing the profile of the size distributions.

For the two reasons described above, it would be desirable to test various ways of encapsulation techniques in order to obtain even smaller water compartments, which might be more stable.

Third, after extraction of the water phase from the emulsion, free M.Hae III fusion proteins, which are not linked to their encoding DNA molecule, might bind to other DNA molecules in a non-covalent manner and, therefore, lead to the selection of DNA molecules coding for non-binding proteins. Hae III DNA-methyltransferase binds its substrate sequence 5′-GGCC-3′ with nanomolar affinity (Chen *et al*., 1993). In order to capture free M.Hae III fusion proteins, short DNA

fragments (1 μM) containing five times the DNA-methyltransferase substrate sequence 5′-GGCC-3′ were added during extraction to the water phase. However, the high affinity of M.Hae III for its substrate sequence is advantageous and a prerequisite for a fast and efficient cross-linking reaction because one DNA molecule in a water compartment with a diameter of 1 μm corresponds to a concentration of 3 nM. Different fusion constructs of Hae III DNA-methyltransferase were expressed *in vitro* with high yields and without signs of proteolysis (Figure 5), which promotes M.Hae III as a good fusion partner for many different proteins.

The results obtained by model selection experiments indicate the robustness and efficiency of this selection system, which may be useful for the directed evolution of peptides and small globular proteins for the generation of novel protein therapeutics and bioseparation reagents (Tomlinson, 2004). Owing to the covalent link between protein and DNA, this selection methodology could be especially well suited for the engineering of proteins which are required to be stable under harsh conditions, such as high temperatures, presence of denaturing agents, extremes of pH, high or low salt concentrations and organic solvents. Such proteins could find various applications in many different fields of industrial biotechnology and nanotechnology.

Recently, the work of two other groups has been published in the field of DNA display. The first group used the strong noncovalent interaction of the biotin-streptavidin complex to couple proteins with DNA (Yonezawa *et al*., 2003). A biotinylated DNA library coding for random decapeptides fused to streptavidin was expressed in the water compartments of a water-in-oil emulsion and peptides binding to anti-Flag antibody were successfully selected from the resulting library of DNA–protein complexes. The strength of the interaction between biotin and streptavidin is advantageous. However, in a previous study by the same group (Doi and Yanagawa, 1999), only 1% of input DNA associated with peptides, probably owing to low yields of streptavidin-fused peptides expressed in an *E.coli* S30 *in vitro* transcription/translation system. The situation could be improved by using a wheat germ *in vitro* transcription/translation system leading to the formation of DNA–peptide conjugates with an efficiency of >95%. It remains to be seen whether streptavidin fused to larger proteins can also be expressed with sufficiently high yields.

The second group (Odegrip *et al*., 2004) presented an alternative DNA-based approach which does not require encapsulation for correct assigment of proteins with their encoding DNA molecules. This is possible due to the *cis*-acting activity of RepA protein, which binds with high fidelity to its encoding DNA molecule. However, the interaction of RepA with its binding site on DNA is neither covalent nor as strong as that of biotin with streptavidin, which may impair the selection of binding proteins with slow dissociation kinetics (low k_{off}).

With the approach presented here, a combination of a very stable physical linkage and high expression levels of different fusion proteins was achieved. In addition, the number of M.Hae III fusion proteins linked per DNA molecules might be varied by simply adjusting the number of M.Hae III substrate sequences containing the modified base 5-fluorodeoxycytidine in the DNA molecule. Covalent complexes of a DNA gene with a defined, small number of copies of the encoded protein could also reduce the bias in selection due to different expression levels. Coupling proteins covalently to DNA is potentially not only useful for selection of binding activity, but DNA–protein fusions could also be re-compartmentalized in emulsions and selected for catalysis (Griffiths and Tawfik, 2003). In addition, more complex selection protocols for the engineering of allosteric proteins could also be applied more easily owing to the covalent nature of the linkage between protein and DNA.

Acknowledgements

The financial support of Novartis (bursary to J.B.) and of the Gebert-Rüf Foundation is gratefully acknowledged.

References

Bertschinger,J. and Heinis,C. (2003) *Swiss Patent Application*, 01671/03.
Chen,L., MacMillan,A.M., Chang,W., Ezaz-Nipkay,K., Lane,W.S. and Verdine,G.L. (1991) *Biochemistry*, **30**, 11018–11025.
Chen,L., MacMillan,A.M., Verdine,G.L. (1993) *J. Am. Chem. Soc.*, **115**, 5318–5319.
Doi,N. and Yanagawa,H. (1999) *FEBS Lett.*, **457**, 227–230.
Feldhaus,M.J. *et al*. (2003) *Nat. Biotechnol.*, **21**, 163–170.
Griffiths,A.D. and Tawfik,D.S. (2003) *EMBO J.*, **22**, 24–35.
Li,E., Pedraza,A., Bestagno,M., Mancardi,S., Sanchez,R. and Burrone,O. (1997) *Protein Eng.*, **10**, 731–736.
Mattheakis,L.C., Bhatt,R.R. and Dower,W.J. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 9022–9026.
Nametkin,S.N., Kolosov,M.I., Ovodov,S.Y., Alexandrov,A.N., Levashov,A.V., Alakhov,V.Y. and Kabanov A.V. (1992) *FEBS Lett.*, **309**, 330–332.
Montigiani,S., Neri,G., Neri,P. and Neri,D. (1996) *J. Mol. Biol.*, **258**, 6–13.
Neri,D., de Lalla,C., Petrul,H., Neri,P. and Winter,G. (1995) *Biotechnology*, **13**, 373–377.
Odegrip,R., Coomber,D., Eldridge,B., Hederer,R., Kuhlmann,P.A., Ullman,C., FitzGerald,K. and McGregor,D. (2004) *Proc. Natl Acad. Sci. USA*, **101**, 2806–2810.
Olsen,M.J., Stephens,D., Griffiths,D., Daugherty,P., Georgiou,G. and Iverson,B.L. (2000) *Nat. Biotechnol.*, **18**, 1071–1074.
Roberts,R.W. and Szostak,J.W. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 12297–12302.
Schaffitzel,C., Hanes,J., Jermutus, L. and Plückthun,A. (1999) *J. Immunol. Methods*, **231**, 119–135.
Sepp,A., Tawfik,D.S. and Griffiths,A.D. (2002) *FEBS Lett.*, **532**, 455–458.
Smith,G. (1998) *Trends Biochem. Sci.*, **23**, 457–460.
Tabuchi,I., Soramoto,S., Nemoto,N. and Husimi,Y. (2001) *FEBS Lett.*, **508**, 309–312.
Tawfik,D.S. and Griffiths,A.D. (1998) *Nat. Biotechnol.*, **16**, 652–656.
Tomlinson,I.M. (2004) *Nat. Biotechnol.*, **22**, 521–522.
Walsh,G. (2000) *Nat. Biotechnol.*, **18**, 831–833.
Walsh,G. (2003) *Nat. Biotechnol.*, **21**, 865–870.
Winter,G., Griffiths,A.D., Hawkins,R.E. and Hoogenboom,H.R. (1994) *Annu. Rev. Immunol.*, **12**, 433–455.
Xu,L. *et al*. (2002) *Chem. Biol.*, **9**, 933–942.
Yonezawa,M., Doi,N., Kawahashi,Y., Higashinakagawa,T. and Yanagawa,H. (2003) *Nucleic Acids Res.*, **31**, e118.
Zardi,L., Carnemolla,B., Siri,A., Petersen,T.E., Paolella,G., Sebastio,G. and Baralle,F.E. (1987) *EMBO J.*, **6**, 2337–2342.

Received October 4, 2004; accepted October 23, 2004

Edited by Andrew Griffiths